

iTrans: A computational tool to trace the flow of knowledge from basic research to clinical applications

B. Ian Hutchins^{1†} and Ling Bai¹

¹Office of Portfolio Analysis, DPCPSI, OD [†]bruce.hutchins@nih.gov

Introduction

Translating basic discoveries into clinical applications is a key priority for NIH. Many basic research programs are initiated with the goal of generating new knowledge that will eventually be used in clinical research. We need ways of tracking the knowledge generated by basic research as it moves outside its original field toward new clinical applications.

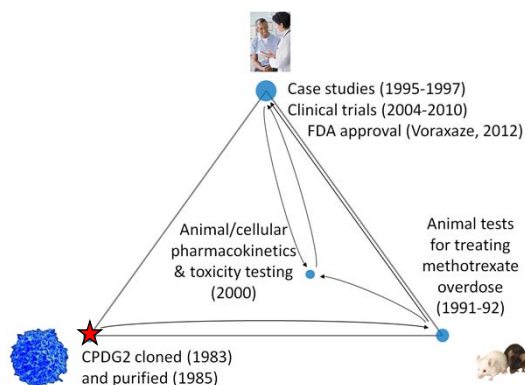
Methods

A recent study¹ published a new algorithm to automatically classify Medline-indexed articles as human, animal, or cellular/molecular research (or some combination of these), based on the location of their MeSH terms² within the MeSH tree. We have further developed and refined this algorithm into a prototype tool for NIH use. The position of a paper or set of papers is mapped onto a triangle, indicating its place within the spectrum of cellular/molecular, animal, and human research. By mapping the positions of subsequent articles that cite the seminal work on this triangle, the movement of knowledge can be traced as it flows from basic knowledge to clinical application.

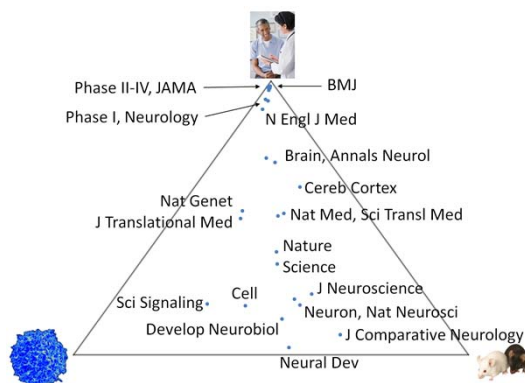
¹Weber G (2013) Identifying translational science within the triangle of biomedicine. J Trans Med 11:126

²Medical Subject Headings (MeSH) is the National Library of Medicine's controlled vocabulary thesaurus. It consists of sets of terms naming descriptors in a hierarchical structure that permits searching at various levels of specificity.

Visualizing the Translation of Knowledge

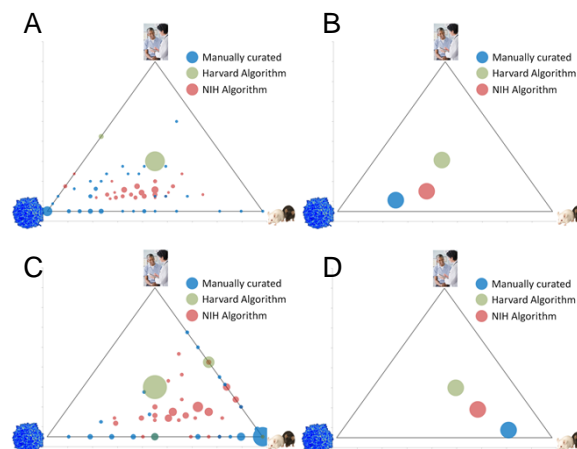


Example of basic research discoveries (★, bottom left) translating to FDA approval through therapeutic development, animal testing, and clinical studies.



Classifications of journals spanning the entire range of human, animal and cellular research, as well as Phase I-IV clinical trials.

Hard Cases: Worst-Case Scenario Performance

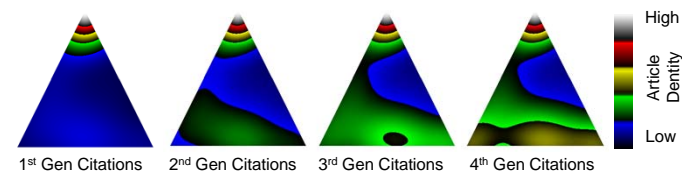


- (Top) Attempts to confuse the algorithm with ambiguous terms (e.g. searching PubMed for "human embryonic kidney [Hek293] cells", which is cell biology research, not human research). (A) Bubble plots depict positions of papers (bubble size ~ number of papers) determined by manual curation (light blue), the original Harvard algorithm¹ (green) or the new NIH algorithm (red). (B) The average of all papers for each of the three methods.
- (Bottom) The same experiment using the search term "humanized mouse model." (C) Positions of individual papers for the three different plotting methods, (D) the average over all papers for each of the three methods.

PubMed queries for the hard cases were:

(A,B) Hek293 cells AND breast cancer AND mouse ("1990"[Date - Completion] : "2012"[Date - Completion])
(C,D) (humanized mouse model) AND ("1990"[Date - Completion] : "2012"[Date - Completion])

Screening for Basic Research Contributions



iTrans can be used to screen for basic research that culminates in a clinical application; in the example shown here, citations were traced backward in time (triangles, left to right) from clinical trials published in 2012. Articles directly cited by clinical trials were overwhelmingly human research (1st Gen citations, left), but tracing back further reveals a chain of articles with progressively higher levels of basic research. Only articles published within 10 years of the trial were considered (>40,000 clinical trials and >2,000,000 cited articles).